

Physics 521, Quantum I

Lecture 5-6: The Structure of Quantum Mechanics

With the mathematical tools of Hilbert space and linear operators in hand, we are now prepared to lay out the structure of quantum theory. To get started we begin with the "postulates" as laid out in the so called Copenhagen interpretation, spearheaded by Niels Bohr in the late 1920's. The Copenhagen interpretation remains the foundation of quantum theory to this day, though with some important modifications, as we will describe.

The "Copenhagen Interpretation": A set of postulates that describe an operational procedure for predicting the outcome of experiments

- The State: The state of a system is a mathematical tool for calculating the probability of outcomes of experiments. If we have maximal possible information about the system, we represent it as a vector in Hilbert space

$$|\psi\rangle \equiv \text{State} \in \mathcal{H}$$

Notes:

- Such states are called "pure states," to be contrasted with "mixed states" when we have incomplete information, as we will see.

- $|\psi\rangle$ and $\alpha|\psi\rangle$ are considered to be the same state for any complex scalar α , i.e., pure states are "rays" in Hilbert space. We typically restrict states to be "normalized" as $\langle\psi|\psi\rangle=1$. The state is still arbitrary up to an overall phase ϕ : $|\psi\rangle \equiv e^{i\phi}|\psi\rangle$

- Relative phases in an expansion in a basis are not negligible.

$$\begin{aligned} \text{E.g. Spin-}\frac{1}{2} \quad |\psi\rangle &= \alpha|+_z\rangle + \beta|-_z\rangle = |\alpha|e^{i\phi_\alpha}|+_z\rangle + |\beta|e^{i\phi_\beta}|-_z\rangle \\ &= |\alpha||+_z\rangle + |\beta|e^{i(\phi_\alpha - \phi_\beta)}|-_z\rangle \quad \phi = \phi_\alpha - \phi_\beta \end{aligned}$$

Normalization constrains $|\alpha|^2 + |\beta|^2 = 1 \Rightarrow$ Pure state specified by 2 real parameters

Generally, in d -dimensions, # of real parameters = $2d - 2$

2 complex amplitudes - 1 for phase - 1 for norm

- Observables: Every Hermitian operator is associated with a "physical observable," i.e. something that can be measured. This includes familiar physical quantities from classical physics such as position, momentum, energy, orbital angular momentum.

Notes: • This postulate is often stated in reverse: observables are Hermitian operators. This is not strictly true, e.g. there is no Hermitian operator for time. Any there are measurements we can do which you might say correspond to nonhermitian operators (as we will see)

- There are nonclassical observables, such as "spin angular momentum"

- Measurement Outcomes: The eigenvalues of a Hermitian operator \hat{A} correspond to possible "fine grained" outcomes one can find in a measurement of \hat{A} .

Note: Most measurement apparatuses don't yield such "sharp" measurements, as we will see.

- Probability of measurement outcome: The probability of finding a given measurement outcome is given by the "Born rule"

(i) Suppose $|a\rangle$ is a nondegenerate eigenvector of \hat{A} with eigenvalue a
 \Rightarrow Probability of finding "a" in a measurement of \hat{A} is

$$P_a = |\langle a | \psi \rangle|^2 \quad (\text{assuming } |a\rangle \text{ and } |\psi\rangle \text{ are normalized})$$

Remember $|\psi\rangle = \sum_a |a\rangle \langle a | \psi \rangle = \sum_a c_a |a\rangle \Rightarrow P_a = |c_a|^2$ (square of prob. amplitude)

Note: Because the state is normalized, $\langle \psi | \psi \rangle = \sum_a |c_a|^2 = \sum_a P_a = 1$
: Total probability for some outcome = 1.

(ii) If \hat{A} has degenerate eigenvectors $\{|u_a^{(i)}\rangle\}$, $\hat{A}|u_a^{(i)}\rangle = a|u_a^{(i)}\rangle$: $i=1, 2, \dots, g_a^{\text{degeneracy}}$
Then we must treat different situations differently. If the apparatus

does not determine the other eigenvalues in a complete set of commuting operators that uniquely define the state, then we must add probabilities within the degenerate subspace:

$$P_a = \sum_{i=1}^{g_a} |\langle u_a^{(i)} | \psi \rangle|^2$$

If the apparatus also measures the compatible observables in the mutually commuting set, then $P_{a,b,c\dots} = k_{a,b,c\dots} |\psi\rangle^2$

Note: This application of the Born rule is for "projective measurements"

- Nondegenerate case: 1D projector $\hat{P}_a = |a\rangle\langle a|$

$$\text{probability } P_a = k_a |\psi\rangle^2 = \|\hat{P}_a |\psi\rangle\|^2 = \langle \psi | \hat{P}_a \hat{P}_a |\psi \rangle = \langle \psi | \hat{P}_a |\psi \rangle$$

- Degenerate case: $\dim g_a^{(i)}$ projector $\hat{P}_a = \sum_{i=1}^{g_a^{(i)}} |u_a^{(i)}\rangle\langle u_a^{(i)}|$

$$\text{probability } P_a = \sum_i k_{u_a^{(i)}} |\psi\rangle^2 = \sum_i \langle \psi | u_a^{(i)} \rangle \langle u_a^{(i)} | \psi \rangle = \langle \psi | \hat{P}_a |\psi \rangle$$

The resolution of the identity ensures $\sum_a \hat{P}_a = \hat{1} \Rightarrow$ Probabilities sum to 1.

Note: Not all measurements are "projective." To specify probabilities of measurement outcomes we require a resolution of the identity in terms of positive operators $\{\hat{E}_a\}$, i.e. all operators $\hat{E}_a \geq 0$ (no negative eigenvalues).

If $\sum_a \hat{E}_a = \hat{1}$, then $P_a = \langle \psi | \hat{E}_a | \psi \rangle$. This most general type of measurement is known as a **POVM** (which has a heavy handed name: positive operator valued measure)

The important fact of nature is that measurement outcomes cannot be predicted definitively, even for a perfect measurement apparatus with no noise. The outcome is random and we can only assign probabilities to outcomes. Some key statistical properties have uniquely quantum mechanical expressions

Expected Value: Given a random variable "a", with probability P_a

$$\bar{a} = \sum_a a P_a$$

We $\{a\}$ are the eigenvalues of \hat{A} , and $P_a = \langle \psi | \hat{P}_a | \psi \rangle$

$$\bar{a} = \sum_a a \langle \psi | \hat{P}_a | \psi \rangle = \langle \psi | \underbrace{\sum_a a \hat{P}_a}_{\hat{A}} | \psi \rangle = \langle \psi | \hat{A} | \psi \rangle = \langle \hat{A} \rangle_{\psi}$$

(expectation value of \hat{A} in $|\psi\rangle$)

Uncertainty: ΔA

$$\text{Variance: } \Delta A^2 = \langle \Delta \hat{A}^2 \rangle \equiv \langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle$$

$$\Rightarrow \Delta A^2 = \langle \hat{A}^2 \rangle - 2\langle \hat{A} \rangle^2 + \langle \hat{A} \rangle^2 \Rightarrow \boxed{\Delta A^2 = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2}$$

$$\text{Uncertainty} = \text{rms } \Delta A = \sqrt{\langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2}$$

Note $\Delta A = 0 \Rightarrow |\psi\rangle$ is an eigenstate of \hat{A}

Heisenberg Uncertainty Relation and Incompatible Observables

Suppose we have two observables that don't commute $[\hat{A}, \hat{B}] \neq 0$

$\Rightarrow \hat{A}$ and \hat{B} do not share complete set of eigenvectors \Rightarrow In general we cannot simultaneously predict the outcome of $\hat{A} + \hat{B}$ with certainty

Formally: $\boxed{\Delta A \Delta B \geq \frac{1}{2} |\langle [\hat{A}, \hat{B}] \rangle|}$

Proof: Let $|\phi_A\rangle = \frac{\Delta \hat{A}}{\Delta \hat{B}} |\psi\rangle = \frac{(\hat{A} - \langle \hat{A} \rangle)}{(\hat{B} - \langle \hat{B} \rangle)} |\psi\rangle$ (Unnormalized vectors)

$$\Rightarrow \langle \Delta \hat{A}^2 \rangle = \langle \phi_A | \phi_A \rangle, \quad \langle \Delta \hat{B}^2 \rangle = \langle \phi_B | \phi_B \rangle$$

Now, according to the Schwartz inequality: $\langle \phi_A | \phi_A \rangle \langle \phi_B | \phi_B \rangle \geq |\langle \phi_A | \phi_B \rangle|^2$

$$\Rightarrow \Delta A^2 \Delta B^2 \geq |\langle \psi | (\hat{A} - \langle \hat{A} \rangle)(\hat{B} - \langle \hat{B} \rangle) | \psi \rangle|^2 = |\langle \psi | (\underbrace{\frac{1}{2} [\Delta \hat{A}, \Delta \hat{B}]}_{\text{imaginary}} + \underbrace{\frac{1}{2} \{ \Delta \hat{A}, \Delta \hat{B} \}}_{\text{real}}) | \psi \rangle|^2$$

$$\Rightarrow \Delta A^2 \Delta B^2 \geq \frac{1}{4} |\langle \psi | [\Delta \hat{A}, \Delta \hat{B}] | \psi \rangle|^2 + \frac{1}{4} |\langle \psi | \{ \Delta \hat{A}, \Delta \hat{B} \} | \psi \rangle|^2 \geq \frac{1}{4} |\langle \psi | [\Delta \hat{A}, \Delta \hat{B}] | \psi \rangle|^2 = \frac{1}{4} |\langle \psi | [\hat{A}, \hat{B}] | \psi \rangle|^2$$

q.e.d

Dynamics: How does the state evolve as a function of time? Quantum theory has two distinctly different types of evolution:

- Deterministic time evolution: If we do not learn anything about the state of the system, then there is a determined map that takes the state at one time to the state at a later time. If the system is closed, i.e., so "stuff" passes into or out of the system, then we neither gain nor lose information about the system

\Rightarrow In a closed system, a pure state remains a pure state:

$$\underbrace{\hat{M}(t+t_0, t_0)}_{\text{map } t_0 \rightarrow t+t_0} |\psi(t_0)\rangle = |\psi(t+t_0)\rangle$$

Since we neither gain nor lose information, the total probability of finding measurement outcomes is conserved $\Rightarrow |||\psi(t+t_0)\rangle|| = 1$

$$\Rightarrow \langle \psi(t+t_0) | \psi(t+t_0) \rangle = \langle \psi(t_0) | \hat{M}^\dagger(t+t_0, t_0) \hat{M}(t+t_0, t_0) | \psi(t_0) \rangle = 1$$

$\Rightarrow \hat{M}(t+t_0, t_0)$ must be unitary.

\Rightarrow An isolated system evolves in time according to a unitary map

$$|\psi(t+t_0)\rangle = \hat{U}(t+t_0, t_0) |\psi(t_0)\rangle$$

Note: this is reversible, i.e. $|\psi(t_0)\rangle = \hat{U}^\dagger(t+t_0, t) |\psi(t+t_0)\rangle$

- Nondeterministic evolution (Measurement)

We now come to the most confusing and contentious part of quantum theory, the "measure postulate," sometimes known as the "collapse of the wave function." When we "measure" a system, e.g. we do a fine grained measurement of an observable \hat{A} , and we find a eigenvalue a , immediately after the measurement, the state "collapses" onto the corresponding eigenvector

Collapse: "Measure \hat{A} ", find (nondegenerate) eigenvalue $|a\rangle$

$$|\psi\rangle = \sum_a c_a |a\rangle \Rightarrow |a\rangle \text{ with probability } |c_a|^2$$

The "collapse of the wavefunction" is distinctly different from the unitary evolution of an isolated system: it is irreversible and stochastic (evolution is not deterministic)

The update rule on the state, conditional on the this measurement outcome can formally be written:

- Given $|\Psi(t_s)\rangle$ (t_s = time just before measurement)
- Measurement result a .
- State immediately after the measurement:

$$|\Psi(t_f)\rangle = \frac{\hat{P}_a |\Psi(t_s)\rangle}{\sqrt{\hat{P}_a |\Psi(t_s)\rangle}} = \frac{\hat{P}_a |\Psi(t_s)\rangle}{\sqrt{\langle \Psi(t_s) | \hat{P}_a | \Psi(t_s) \rangle}} = \frac{\hat{P}_a |\Psi(t_s)\rangle}{\sqrt{p_a}}$$

- If "a" is a nondegenerate eigenvalue, $\hat{P}_a = |a\rangle \langle a|$, $p_a = |\langle a | \psi \rangle|^2$
 $\Rightarrow |\Psi(t_f)\rangle = \frac{|a\rangle \langle a | \Psi(t_s)\rangle}{|\langle a | \Psi(t_s)\rangle|} = |a\rangle$ up to an overall phase
 $|\Psi\rangle$ is projected onto the eigenstate corresponding to the eigenvalue
- If "a" is degenerate, and the device cannot in any way distinguish the other compatible observables that commute with \hat{A} , the $|\Psi\rangle$ is projected onto the subspace H_a .

The "resetting" of the state $|\Psi\rangle$, generally a superposition $|\Psi\rangle = \sum_a c_a |a\rangle$, to the state $|a\rangle$ is often known as the collapse of wavefunction. No other quantum effect has continued to mystify and confuse us as collapse.

Some mysteries of "collapse"

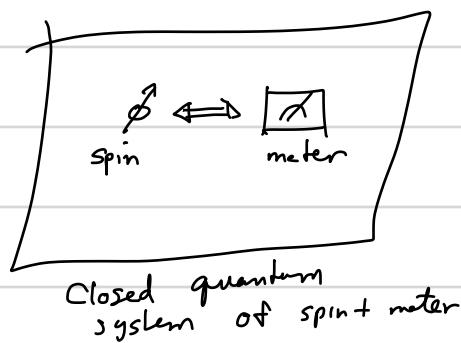
We have seen that quantum theory has two distinctly different types of evolution

(i) Deterministic, continuous $|\Psi(t+t_0)\rangle = \underbrace{U(t+t_0, t)}_{\text{smooth function of } t} |\Psi(t_0)\rangle$

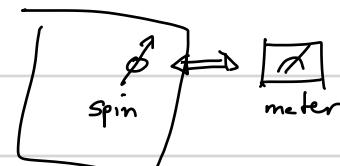
(ii) Random, "quantum jump" $|\Psi(t)\rangle \Rightarrow$ (a) with probability $p_a = |\langle a | \Psi(t) \rangle|^2$

Under what physical circumstances does the state evolve according to (i) vs. (ii)?

The evolution (ii) is the rule when we do a "measurement" on the system. But which physical interactions constitute measurement, and which are "closed quantum system" of many degrees of freedom.



vs.



Closed quantum system of spin measured by "meter" outside

Bohr (and the Copenhagen interpretation) answered this problem by dividing the world into the "quantum world" and the "classical world". A measurement occurs when a quantum system is observed by a "classical measuring device". We can see that this is somewhat undefined, ad hoc, and perhaps even inconsistent. At what point is a measuring device considered "classical" and thus has the capacity to effect a collapse of the wavefunction? And if all physical phenomena are fundamental quantum, how can a measuring device which thus must arise from such quantum phenomena act outside the law of quantum mechanics.

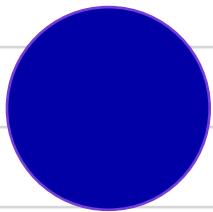
The search for the classical/quantum boundary has been one of the great pursuits in explorations of the foundations of quantum mechanics. But clearly, this boundary cannot be sharp and whereas we should be able to see classical physics "emerge" as an effective limiting case of quantum mechanics (similar to how Galilean relativity follows as a limit of Einstein special relativity), we can limit from deterministic to stochastic evolution.

What is collapsing?

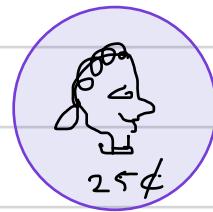
Is the collapse of physical process? This goes to a question we have so far not completely addressed. What is the state $|ψ\rangle$? Is it a physical thing, like a wave on a string, or even something as abstract as a wave of the electromagnetic field? If $|ψ\rangle$ is a physical "thing" then the collapse must be a physical process, and we should be able to describe the physical mechanism by which it occurs.

But what is the alternative? If $|ψ\rangle$ is a "physical thing" existing in the world, what is it? A modern perspective is that the quantum state is a state of knowledge, i.e., it is a description above what we know about the system. When we learn something, we update our state of knowledge. This is known as the Bayesian perspective, following the Bayesian theory of probability. For a Bayesian, probabilities are "degrees of belief" (i.e., subjective entities, rather than objective facts about the world). But, one acts as a rational player. So, one should update one probability assignment when one gains new information.

For example, suppose we have a coin, but it is covered. Since you don't know anything, you would assign a 50-50 probability to the coin being face-up heads or face-up tails. But if the coin is uncovered, and we learn it is heads, then we would update a probability assignment to $P_{\text{heads}} = 1$, $P_{\text{tails}} = 0$



$$P_{\text{heads}} = P_{\text{tail}} = \frac{1}{2}$$



$$P_{\text{heads}} = 1, P_{\text{tail}} = 0$$

More generally, according to Bayes' Rule, if we assign a "prior probability" $P(x)$ to be true, then if we learn that y is true we must update to the "posterior probability"

$$P(x|y) = \frac{P(y|x) P(x)}{P(y)}$$

↑
Prob of x given y

In Quantum theory, if the state of the state is $|\Psi_{in}\rangle$, we assign probability of finding "x" in a projective measurement according to the Born rule: $P(x) = \langle \Psi_{in} | \hat{P}_x | \Psi_{in} \rangle$. If we do a projective measurement and find "y", then we must update the state as

$$(\text{output state given we find } y) \quad |\Psi_{out}\rangle_y = \frac{\hat{P}_y |\Psi_{in}\rangle}{\|\hat{P}_y |\Psi_{in}\rangle\|}$$

$$\Rightarrow P_{\text{posterior}}(x|y) = \langle \Psi_{out} | \hat{P}_x | \Psi_{out} \rangle_y = \frac{\langle \Psi_{in} | \hat{P}_x \hat{P}_y \hat{P}_x | \Psi_{in} \rangle}{\langle \Psi_{in} | \hat{P}_y | \Psi_{in} \rangle} \leftarrow P_{\text{prior}}(y)$$

This is a form of Bayes rule, restricted to projective measurements. We will see that it looks just like the classical Bayes rule for more general POVMs on compatible measurement

So, is the collapse of the wave function just Bayesian updating of classical probability theory? This isn't the whole story; quantum mechanics has features completely different from classical probability.

• Quantum "back action" on incompatible observables:

Suppose $[\hat{A}, \hat{B}] \neq 0$ (they are "incompatible observables"). Suppose $|b\rangle$ is an eigenstate of \hat{B} , $\hat{B}|b\rangle = b|b\rangle$, $B|b\rangle = b|b\rangle$. Our uncertainty of the value of \hat{B} we will find is zero $\Delta B = 0$. Now because $[\hat{A}, \hat{B}] \neq 0$, $|b\rangle$ is not generally an eigenstate of \hat{A} as well $\Rightarrow |b\rangle = \sum_a c_a |a\rangle$. If we now perform a projective measurement of \hat{A} we will randomly find "a" with probability $p_a = |c_a|^2$. Our uncertainty of \hat{A} is now zero, $\Delta A = 0$, but we have lost information about \hat{B} . That is, if we now measure \hat{B} we get a random result, and we can only give the probability of an outcome. This is completely nonclassical phenomenon. Gaining information about one observable should not cause us to lose information about another. This is the heart of the uncertainty principle.

• Local realism

In the example of the coin, it's clear what we mean by probability and "state of knowledge". When we assign a probability different from 1 or 0 to heads vs. tails, it is because we are missing complete information. But we believe the information is there, under the cover.

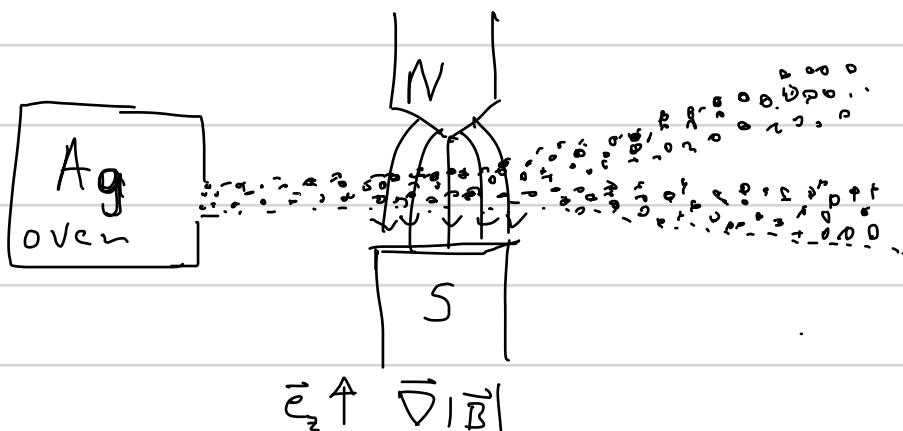
We just don't know it. When we uncover the coin, we learn what was hidden from us and "complete of knowledge." But, as well will see, the quantum world is different. When we do a measurement or find an outcome, we cannot say that the value was always "here," locally to system, and we just discovered what was "hidden" from us. Quantum mechanics is incompatible with local hidden variables. This is one of the strangest results, found leaves us grasping for straws.

If quantum states are "states of knowledge," knowledge about what? It's not knowledge about a local realistic value. We continue to grapple with these issues. The Bayesian perspective is on route - with its roots here at UNM, formulated in the group of Carl Caves, and currently carried forward by his former group members Chris Fuchs + Ruediger Schack along with David Mermin. This School of Thought is known as "qbism" (pronounced cubism), for "quantum Bayesianism."

Regardless, it is clear that quantum theory is a "meta theory" It requires an ingredient outside of itself to say "and now something different happens" and the state collapses. In the standard Copenhagen interpretation of Bohr, it is the classical measuring device that provides the dividing line. The classical measuring device is "meta" - outside the quantum world. In qbism the observer is "meta," our Consciousness is outside the description of science itself.

Example: Stern-Gerlach

One of the most familiar scenarios for implementing a quantum measurement is the Stern-Gerlach apparatus. This device is designed to measure the projection of spin along some direction.



Silver atoms in a collimated beam are sent through a region with an inhomogeneous \vec{B} field whose magnitude has a gradient along an axis (we call \vec{e}_z)

The gradient field exerts a force on a magnetic moment $\vec{\mu}$, according to the law familiar in classical physics $\vec{F} = \vec{\nabla}(\vec{\mu} \cdot \vec{B}) = (\vec{\mu} \cdot \vec{\nabla})\vec{B}$

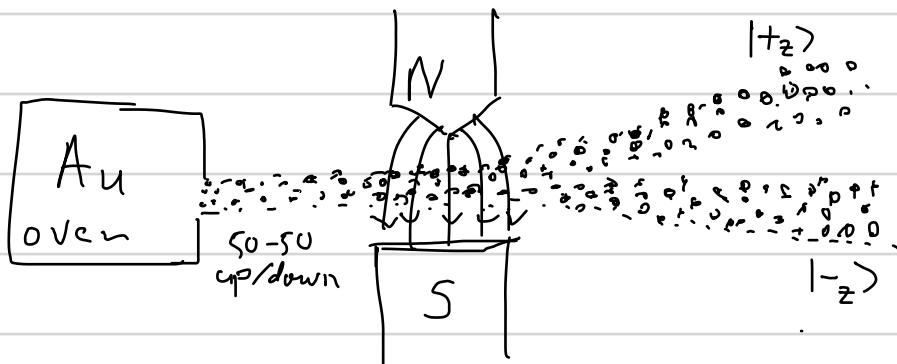
The atom has a magnetic moment due to the spin of the electron:

$$\hat{\vec{\mu}}_P = -2M_B \frac{\hat{\vec{S}}}{\hbar} = -M_B \frac{\hat{\vec{\sigma}}}{\hbar} \quad \text{Bohr magneton } \frac{e\hbar}{2m_e c} \text{ (cgs units)}$$

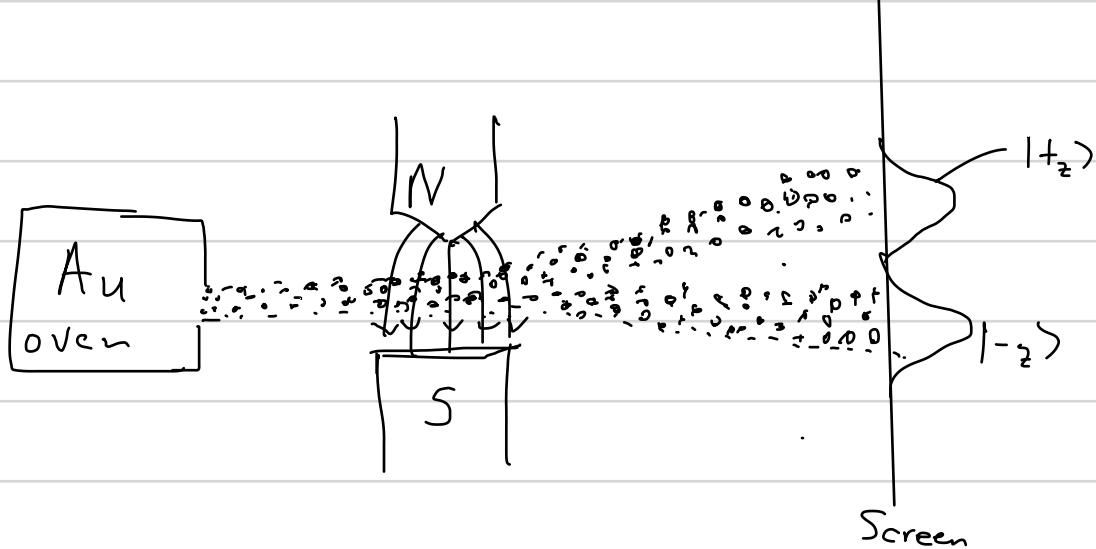
magnetic moment is now an operator (physical quantum observable). Take \vec{B} as a classical field ("c-number")

$$\Rightarrow \hat{\vec{F}} = M_B \frac{\partial |\vec{B}|}{\partial z} \hat{\sigma}_z \vec{e}_z$$

Since the eigenvalues of $\hat{\sigma}_z$ are ± 1 , there are two possible forces. Spin up atoms are deflected up and spin down atoms are deflected down, leading to a separation into two beams

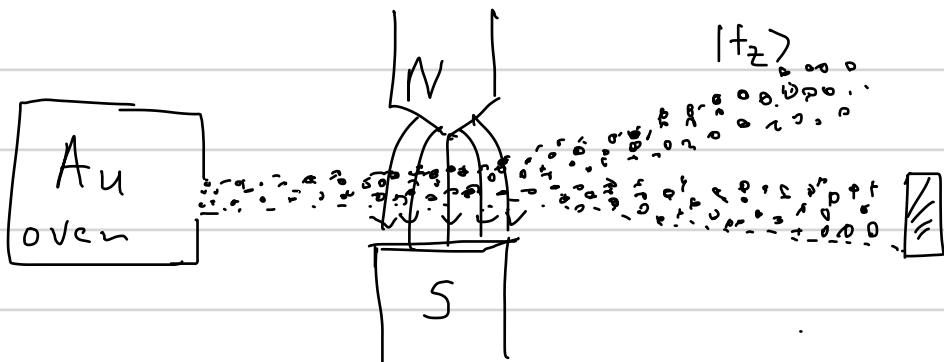


If we put a screen far enough away, we will see two resolved spots, with ~50% in spin up and ~50% in spin down since the spin out of the oven was "random".

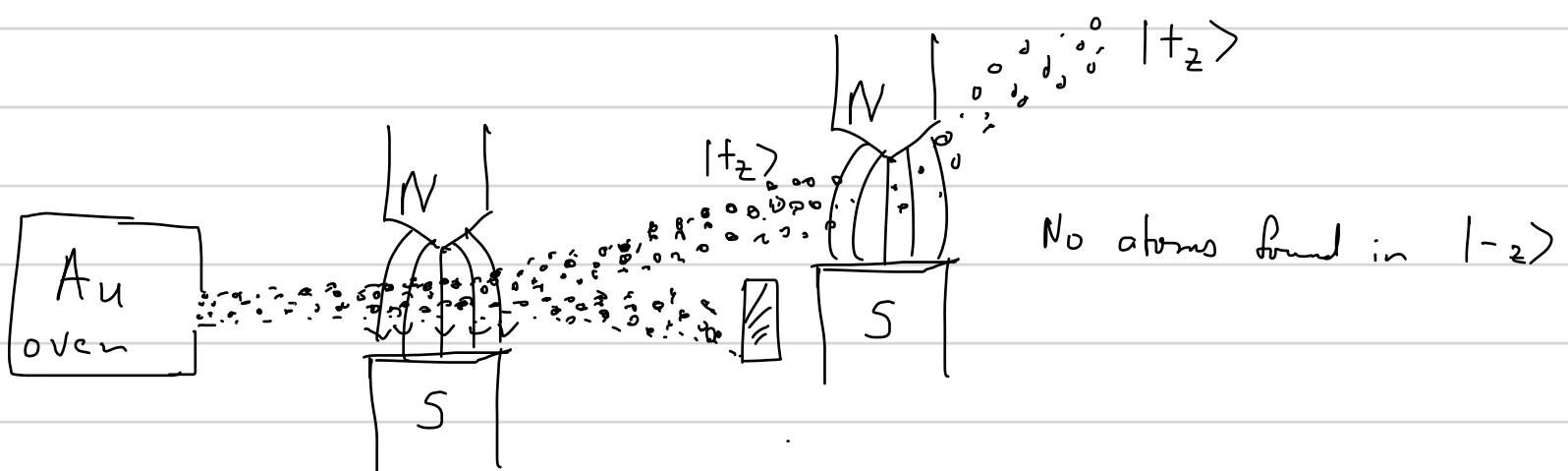


Is this a projective measurement? Well, yes and no. The probability of landing in one spot or the other is $\frac{1}{2}$, but, after the measurement the atoms are absorbed into the screen and chemically react. We can't measure the spin again!

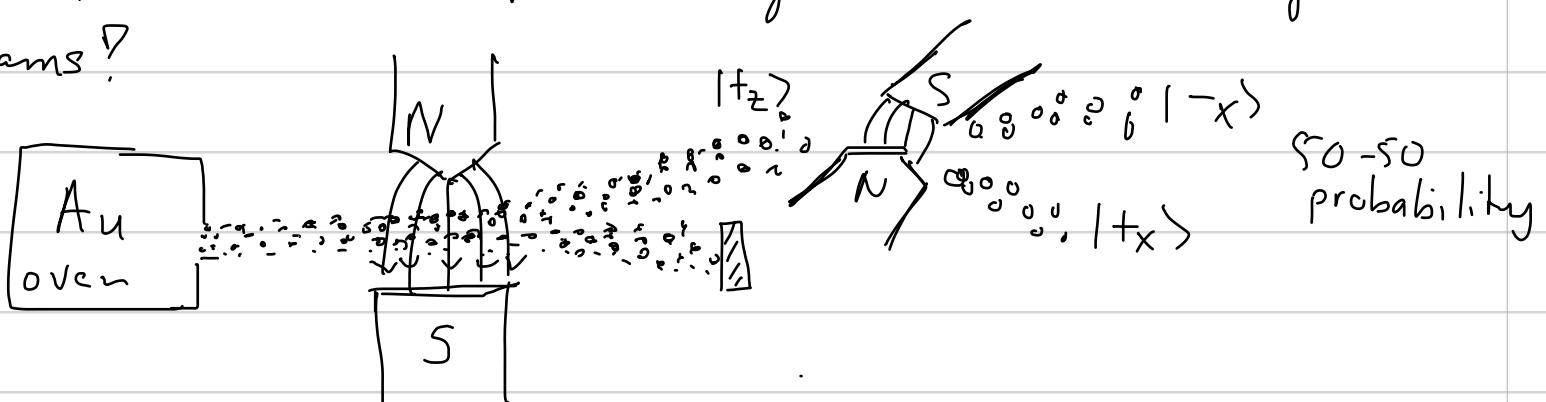
Something closer to a projective measurement: blocking one of the paths



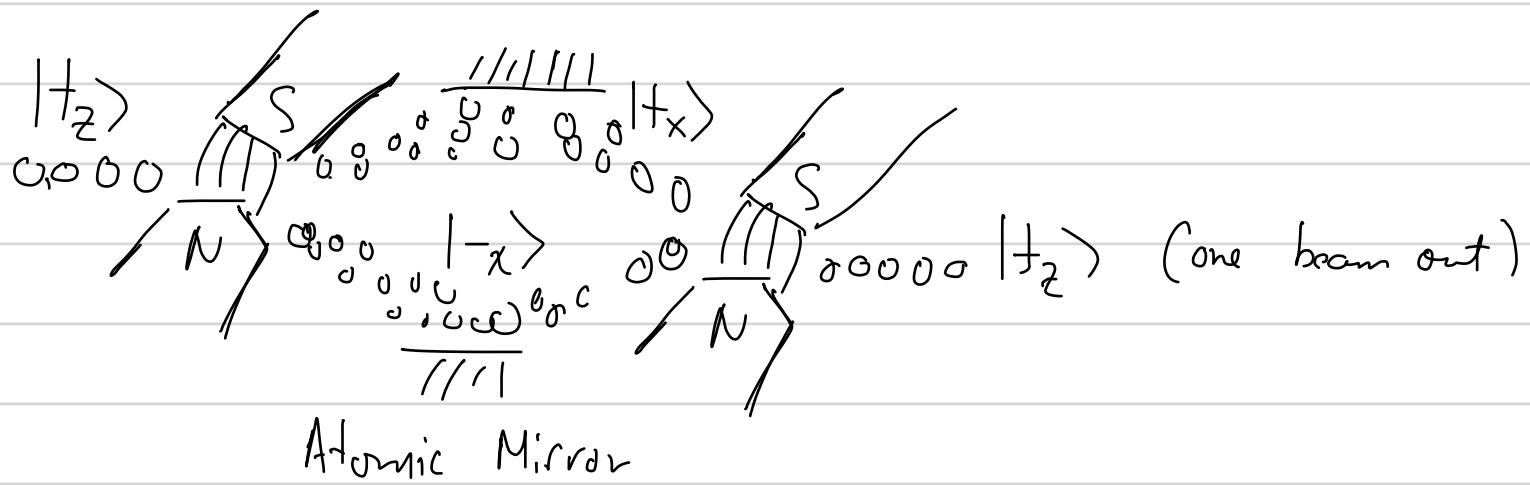
Now, with 50% probability, the atoms in the upper path are definitely in spin-up along z , $|+z\rangle$. Of course the lower path atoms are absorbed, so this is not quite the Bell paradigm. Nonetheless, if we measure the atoms in the upper path again, we are guaranteed to see the same result, $|+z\rangle$; only one beam emerges



On the other hand, if the beam is passed through a SG device along x we see two beams?



Now, did the measure occur? For suppose (with steering magnets) we can recombine the beams again. As discussed in the first lecture, these two beams will interfere!

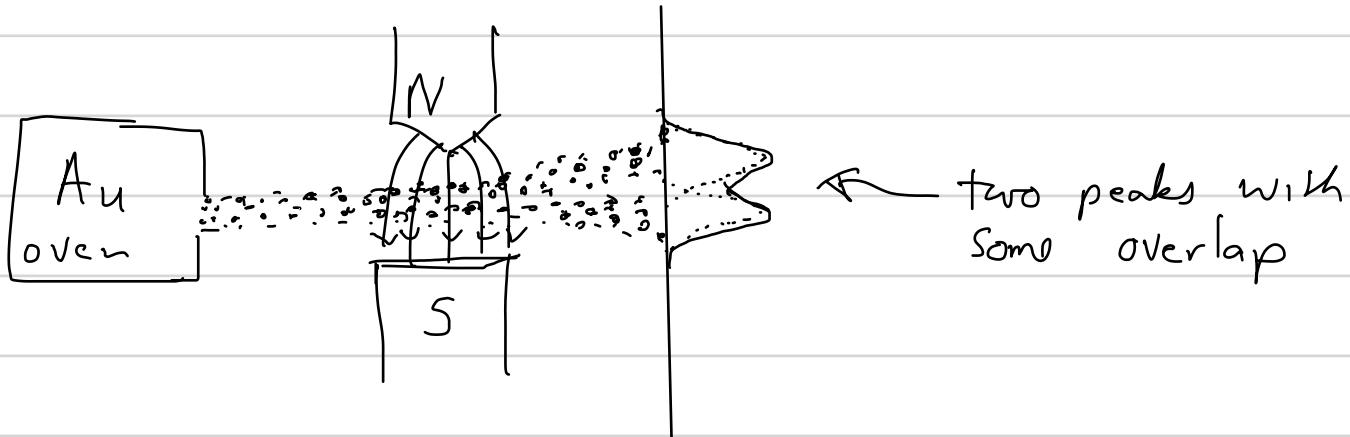


Even though the spins pass through the first SG device, they can (with care) remain in a coherent quantum superposition. This means the measurement was not done.

In a true measurement, the possibility of interference between different alternatives is lost — only one alternative is "realized." This touches on Bohr's concept of the "classical measuring device." One feature of such a device is that "coherence" (i.e. the possibility of seeing interference) between the different alternatives is lost. This process is called decoherence (we will study this in more detail later). This process is closely related to equilibration in classical physics. The microscopic laws of classical physics are reversible but the macroscopic laws are irreversible. Bohr's classical measuring device is "macroscopic" (classical) in the sense that it irreversibly records the measurement outcome. Such irreversible behavior can be seen to arise effectively from microscopic quantum mechanics, similar to the way equilibrium thermal motion can arise effectively in classical mechanics. However, such a theory cannot yield randomness unless it is hidden. The origin of randomness in quantum physics remains the deepest mystery.

Nonprojective measurement

Finally, let's see that the probability of some measurement outcomes are not described by projection operators. Suppose the screen is placed so close to the S-G magnets, that the two beams are not "resolved"



If we wanted to determine whether the atom was $|+z\rangle$ or $| -z\rangle$ we could fit two Gaussians to the measured distribution, and associate one Gaussian \hat{E}_+ and the other \hat{E}_- such that

$$p(\pm) = \langle \psi | \hat{E}_{\pm} | \psi \rangle$$

The operator \hat{E}_+ and \hat{E}_- are not orthogonal — they overlap. Sometimes our assignment \hat{E}_+ with $|+z\rangle$ and \hat{E}_- with $| -z\rangle$ will be incorrect. This is a "fuzzy measurement". In contrast to a projective measurement, the answer can be "maybe."

An example POVM:

$$\hat{E}_+ = 0.95 |+z\rangle \langle +z| + 0.05 |-z\rangle \langle -z|$$

$$\hat{E}_- = 0.05 |+z\rangle \langle +z| + 0.95 |-z\rangle \langle -z|$$

This POVM says that with 95% confidence, if the atom lands in the "+" hump it is spin-up along z , but with 5% chance it was spin down (and vice versa). This is not a projective measurement.